

1 Model

The model takes the form

$$\mathbf{x}_k = \mathbf{f}_k(\mathbf{x}_{k-1}) + \mathbf{v}_k \quad (1a)$$

$$\mathbf{y}_k = \mathbf{H}\mathbf{x}_{k-1} + \mathbf{w}_k \quad (1b)$$

where \mathbf{v}_k and \mathbf{w}_k are white, uncorrelated, zero mean Gaussian noise vectors with covariances of \mathbf{Q} and \mathbf{R} , respectively. For this model the transition density is defined as

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{f}(\mathbf{x}_{k-1}), \mathbf{Q}) \quad (2)$$

$$= (2\pi)^{-d/2} |\mathbf{Q}|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}))^T \mathbf{Q}^{-1} (\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1})) \right) \quad (3)$$

and the likelihood is

$$p(\mathbf{y}_k|\mathbf{x}_k) = \mathcal{N}(\mathbf{H}\mathbf{x}_k, \mathbf{R}) \quad (4)$$

$$= (2\pi)^{-n/2} |\mathbf{R}|^{-1/2} \exp \left(-\frac{1}{2} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k)^T \mathbf{R}^{-1} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k) \right) \quad (5)$$

2 PCRB

The posterior Cramér-Rao bound (PCRB) is computed from the Fisher information matrix computed on the joint density $p(\mathbf{y}, \mathbf{x})$,

$$\mathbf{J} = \mathbb{E} [\nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T \log p(\mathbf{y}, \mathbf{x})] \quad (6)$$

where the expectation is taken over \mathbf{y} and \mathbf{x} . The PCRB bounds the covariance of an unbiased estimator, \mathbf{P} ,

$$\mathbf{P} \geq \mathbf{J}^{-1} \quad (7)$$

Recursive formulas are based on the logarithm of the likelihood and transition densities, which for the model above are

$$\log p(\mathbf{x}_k|\mathbf{x}_{k-1}) = C_1 - \frac{1}{2} (\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}))^T \mathbf{Q}^{-1} (\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1})) \quad (8)$$

$$\log p(\mathbf{y}_k|\mathbf{x}_k) = C_2 - \frac{1}{2} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k)^T \mathbf{R}^{-1} (\mathbf{y}_k - \mathbf{H}\mathbf{x}_k) \quad (9)$$

2.1 Fisher information recursion (non-singular \mathbf{Q})

A recursive formula for the FIM, \mathbf{J}_k , was developed in [?]. The recursion proceeds as

$$\mathbf{J}_k(\mathbf{x}_k) = \mathbf{W}_k - \mathbf{V}_k^T [\mathbf{J}_{k-1}(\mathbf{x}_{k-1}) + \mathbf{U}_k]^{-1} \mathbf{V}_k \quad (10)$$

where

$$\mathbf{U}_k = -\mathbb{E} \left[\nabla_{\mathbf{x}_{k-1}} \nabla_{\mathbf{x}_{k-1}}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) \right] \quad (11)$$

$$\mathbf{V}_k = -\mathbb{E} \left[\nabla_{\mathbf{x}_{k-1}} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) \right] \quad (12)$$

$$\mathbf{W}_k = -\mathbb{E} \left[\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) \right] - \mathbb{E} \left[\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{y}_k | \mathbf{x}_k) \right] \quad (13)$$

The derivatives for the above model are

$$\nabla_{\mathbf{x}_k}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) = -(\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}))^T \mathbf{Q}^{-1} \quad (14)$$

$$\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) = -\mathbf{Q}^{-1} \quad (15)$$

$$\nabla_{\mathbf{x}_{k-1}} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) = \mathbf{F}(\mathbf{x}_{k-1}) \mathbf{Q}^{-1} \quad (16)$$

$$\nabla_{\mathbf{x}_{k-1}}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) = (\mathbf{x}_k - \mathbf{f}(\mathbf{x}_{k-1}))^T \mathbf{Q}^{-1} \mathbf{F}(\mathbf{x}_{k-1}) \quad (17)$$

$$\nabla_{\mathbf{x}_{k-1}} \nabla_{\mathbf{x}_{k-1}}^T \log p(\mathbf{x}_k | \mathbf{x}_{k-1}) = -\mathbf{F}(\mathbf{x}_{k-1})^T \mathbf{Q}^{-1} \mathbf{F}(\mathbf{x}_{k-1}) \quad (18)$$

$$\nabla_{\mathbf{x}_k}^T \log p(\mathbf{y}_k | \mathbf{x}_k) = (\mathbf{y}_k - \mathbf{H} \mathbf{x}_k)^T \mathbf{R}^{-1} \mathbf{H} \quad (19)$$

$$\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{y}_k | \mathbf{x}_k) = -\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (20)$$

Substituting these into (??)-(??) gives,

$$\mathbf{U}_k = \mathbb{E} \left[\mathbf{F}(\mathbf{x}_{k-1})^T \mathbf{Q}^{-1} \mathbf{F}(\mathbf{x}_{k-1}) \right] \quad (21)$$

$$\mathbf{V}_k = \mathbb{E} \left[-\mathbf{F}(\mathbf{x}_{k-1})^T \mathbf{Q}^{-1} \right] \quad (22)$$

$$\mathbf{W}_k = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{Q}^{-1} \quad (23)$$

This method is not applicable when noise is only added to a subset of the states each transition. In this case, \mathbf{Q} is singular and \mathbf{Q}^{-1} does not exist. Another method is required.

2.2 Cramer-Rao bound recursion (singular \mathbf{Q})

An alternative formulation is to recursively compute the bound instead of the matrix, i.e., $\mathbf{P}_k = \mathbf{J}_k^{-1}$ directly rather than \mathbf{J}_k . The formulas suggested in [?] are applicable to wider class of models (e.g. multiplicative noise), denoted by a general transition function $\tilde{\mathbf{f}}(\mathbf{x}_k, \mathbf{v}_k)$. The recursion is

$$\mathbf{P}_{k|k-1} = \bar{\mathbf{F}}_k \mathbf{P}_{k-1|k-1} \bar{\mathbf{F}}_k + \bar{\mathbf{G}}_k \mathbf{Q}_k \bar{\mathbf{G}}_k^{-1} \quad (24a)$$

$$\mathbf{P}_{k|k} = \left(\mathbf{P}_{k|k-1}^{-1} + \bar{\mathbf{R}}_{k-1}^{-1} \right)^{-1} \quad (24b)$$

where

$$\bar{\mathbf{F}}_k = \mathbb{E} \left[\nabla_{\mathbf{x}_k}^T \tilde{\mathbf{f}}_k(\mathbf{x}_k, \mathbf{v}_k) \right] \quad (25)$$

$$\bar{\mathbf{R}}_k^{-1} = -\mathbb{E} \left[\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{y}_k | \mathbf{x}_k) \right] \quad (26)$$

$$\bar{\mathbf{G}}_k = \mathbb{E} \left[\nabla_{\mathbf{v}_k}^T \tilde{\mathbf{f}}_k(\mathbf{x}_k, \mathbf{v}_k) \right] \quad (27)$$

$$\mathbf{Q}_k^{-1} = -\mathbb{E} [\nabla_{\mathbf{v}_k} \nabla_{\mathbf{v}_k}^T \log p(\mathbf{v}_k)] \quad (28)$$

and \mathbf{P}_0^{-1} is the inverse of the prior covariance. The model in (??) is represented by $\tilde{\mathbf{f}}(\mathbf{x}_k, \mathbf{v}_k) = \mathbf{f}(\mathbf{x}_k) + \mathbf{v}_k$. This simplifies the derivatives to

$$\nabla_{\mathbf{x}_k}^T \tilde{\mathbf{f}}_k(\mathbf{x}_k, \mathbf{v}_k) = \mathbf{F}_k(\mathbf{x}_k) \quad (29)$$

$$\nabla_{\mathbf{v}_k}^T \tilde{\mathbf{f}}_k(\mathbf{x}_k, \mathbf{v}_k) = \mathbf{I}_{d \times d} \quad (30)$$

$$\nabla_{\mathbf{x}_k} \nabla_{\mathbf{x}_k}^T \log p(\mathbf{y}_k | \mathbf{x}_k) = -\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (31)$$

and the corresponding matrices to

$$\bar{\mathbf{F}}_k = \mathbb{E} [\mathbf{F}_k(\mathbf{x}_k)] \quad (32)$$

$$\bar{\mathbf{R}}_k^{-1} = \mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} \quad (33)$$

$$\bar{\mathbf{G}}_k = \mathbf{I}_{d \times d} \quad (34)$$

Notice that \mathbf{Q}^{-1} is not directly required in the recursion, so this method is applicable to neural field models where process noise is only added to a single state.

2.3 Monte Carlo approximations

The expectations in (??) and (??) or (??) can be approximated by numerical integration using many trajectory realisations. That is,

$$\mathbf{U}_k \approx \frac{1}{M} \sum_{i=1}^M \mathbf{F}(\mathbf{x}_{k-1}^{(i)})^T \mathbf{Q}^{-1} \mathbf{F}(\mathbf{x}_{k-1}^{(i)}) \quad (35)$$

$$\mathbf{V}_k \approx \frac{1}{M} \sum_{i=1}^M \mathbf{F}(\mathbf{x}_{k-1}^{(i)})^T \mathbf{Q}^{-1} \quad (36)$$

or alternatively,

$$\bar{\mathbf{F}}_{k-1} \approx \frac{1}{M} \sum_{i=1}^M \mathbf{F}(\mathbf{x}_{k-1}^{(i)}) \quad (37)$$

2.4 Analytical evaluation of expectations

For the JR model the expectation in (??) above is

$$\bar{\mathbf{F}}_k = \mathbb{E} [\mathbf{F}_k(\mathbf{x}_k)] \quad (38)$$

$$= \mathbf{F}_k^{(\text{lin})} + \mathbb{E} [\mathbf{G}_k(\mathbf{x}_k)] \quad (39)$$

where $\mathbf{G}_k(\mathbf{x}_k)$ contains the nonlinear component of the Jacobian. These nonlinear terms are derivatives of the error function and have the form,

$$\exp(-(x - v_0)^2 / \zeta^2) \quad (40)$$

The expectation of such terms can only be evaluated when we know the prior mean and covariance at all times.

3 Pendulum example

Let $\mathbf{x}_k = [a_k \ b_k]^T$ where a_k is the pendulum angle and b_k is the angular velocity,

$$\mathbf{f}(\mathbf{x}_k) = \begin{bmatrix} a_k + b_k \Delta \\ -b_k - g \sin(a_k) \Delta \end{bmatrix} \quad (41)$$

The linearisation used for the EKF is

$$\mathbf{F}(\mathbf{x}_k) = \begin{bmatrix} 1 & \Delta \\ -g \cos(a_k) & 1 \end{bmatrix} \quad (42)$$

The observation matrix is $H = [1 \ 0]$, so that only the first state element is measured.

3.1 Simulations

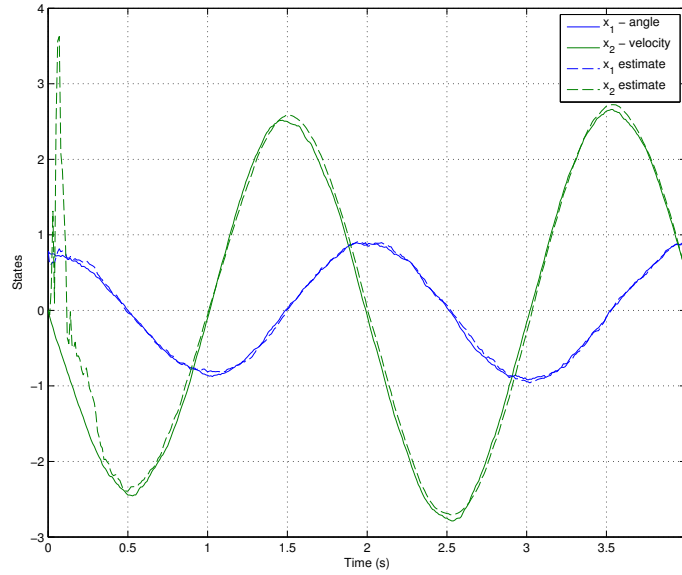


Figure 1: Trajectory of a pendulum and EKF estimates

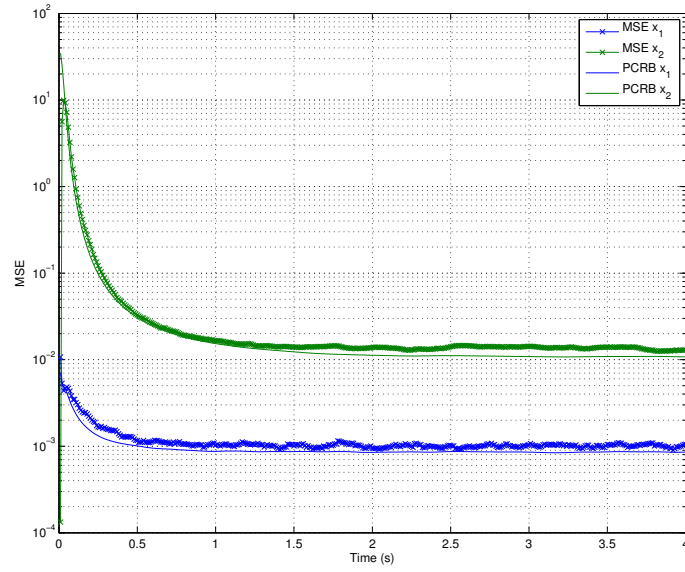


Figure 2: MSE of EKF estimates (averaged over 1000 realisations) compared to the PCRB. The suboptimal EKF is slightly above the PCRB.